

## Stochastic forecasting of COVID-19 daily new cases across countries with a novel hybrid time series model

Arinjita Bhattacharyya<sup>1+</sup>, Tanujit Chakraborty<sup>2+\*</sup>, Shesh N. Rai<sup>1,3,4,5,6</sup>

<sup>1</sup> Department of Bioinformatics & Biostatistics, University of Louisville, KY, USA.

<sup>2</sup> Center for Data Sciences, International Institute of Information Technology, Bangalore, India.

<sup>3</sup> Biostatistics & Bioinformatics Facility, JG Brown Cancer Center, University of Louisville, KY, USA.

<sup>4</sup> The Christina Lee Brown Envirome Institute, University of Louisville, KY, USA.

<sup>5</sup> University of Louisville Alcohol Research Center, University of Louisville, KY, USA.

<sup>6</sup> University of Louisville Hepatobiology & Toxicology Center, University of Louisville, KY, USA.

\* Corresponding author's contact: [tanujit\\_r@isical.ac.in](mailto:tanujit_r@isical.ac.in)

+ These authors have contributed equally.

### ARTICLE HISTORY

Compiled April 25, 2021

### ABSTRACT

1 An unprecedented outbreak of the novel coronavirus (COVID-19) in the form of peculiar pneumonia has spread globally since its first case in Wuhan province, China, in December 2019. Soon after, the infected cases and mortality increased rapidly. The future of the pandemic's progress was uncertain, and thus, predicting it became crucial for public health researchers. These future predictions help the effective allocation of health care resources, stockpiling, and help in strategic planning for clinicians, government authorities, and public health policymakers after understanding the extent of the effect. The main objective of this paper is to develop a hybrid forecasting model that can generate real-time out-of-sample forecasts of COVID-19 outbreaks for five profoundly affected countries, namely the USA, Brazil, India, UK, and Canada. A novel hybrid approach based on the Theta method and Autoregressive neural network (ARNN) model, named Theta-ARNN (TARNN) model, is developed. Daily new cases of COVID-19 are nonlinear, non-stationary, and volatile; thus a single specific model cannot be ideal for future prediction of the pandemic. However, the newly introduced hybrid forecasting model with an acceptable prediction error rate can help healthcare and government for effective planning and resource allocation. The proposed method outperforms traditional univariate and hybrid forecasting models for the test data sets on an average.

### KEYWORDS

20 COVID-19; Forecasting; Theta model; Autoregressive Neural Networks; Hybrid  
21 model; Asymptotic stationarity.  
22

## 23 1. Introduction

24 In December 2019, clusters of pneumonia cases caused by the novel severe acute  
25 respiratory syndrome coronavirus 2 (SARS-Cov-2) were identified at Wuhan, Hubei  
26 province in China [19,22,54]. Soon after the emergence of the novel beta coronavirus,

27 World Health Organization (WHO) characterized this contagious disease as a  
28 pandemic in March 2020 due to its rapid spread within and outside the highly mobile  
29 population of China which got further aggravated by the densely populated location  
30 of the sea-food market and time of Chinese new year [47]. It resulted in an exponential  
31 increase in the incidence rate (IR) and case-fatality rate (CFR) [37]. As of March 16,  
32 2021, a total of 120,512,041 confirmed global cases and 2,665,742 deaths have been  
33 reported worldwide [3].

34  
35 Researchers have faced significant challenges to forecast real-time COVID-19 cases  
36 with traditional mathematical, statistical, and machine learning-based forecasting  
37 tools [9,14,29,31,40,56,60]. Studies in March 2020 with simple and efficient forecasting  
38 methods such as exponential smoothing model predicted cases ten days ahead, with  
39 large confidence intervals, that despite the positive bias, had reasonable forecast error  
40 [43]. Previously used linear and exponential model forecasts for better preparation  
41 regarding hospital beds, ICU admission estimation, resource allocation, emergency  
42 funding, and proposing strong containment measures were conducted [17] that  
43 projected ICU admissions in Italy for March 20, 2020. ICU admission and mechanical  
44 ventilation for critically ill patients reached their peak shattering the health system of  
45 Lombardy, Italy, by end of March 2020 [18]. Health-care workers had to go through the  
46 immense mental stress left with a formidable choice of prioritizing young and healthy  
47 adults over the elderly for allocation of life support, especially unwanted ignoring of  
48 those who are extremely unlikely to survive [13,48]. Real estimates of mortality with  
49 14-day delay demonstrated underestimating the COVID-19 outbreak and indicated  
50 a grave future with a global CFR of 5.7% [2]. Some of the countries are eventually  
51 in a second wave with cases rising in India, and the UK. The uncertain future has  
52 several unanswered questions on the daily trajectory of the pandemic, infection rate,  
53 and number of deaths. Thus, real-time forecasting with time series models could  
54 respond to these debatable quires to reach a statistically validated conjecture in  
55 this current health crisis. Some of the impacting leading-edge research concerning  
56 real-time projections of COVID-19 confirmed cases, recovered cases, and mortality  
57 using statistical, epidemiological and machine learning models are as follows. Short-  
58 term forecasting of cumulative confirmed cases was produced using autoregressive  
59 integrated moving average (ARIMA), Cubist regression, Random Forest, Ridge,  
60 Support Vector Regression (SVR) and Stacking-Ensemble Learning (SEL) model  
61 in [46] whereas Exponential smoothing model produced ten days ahead forecasts of  
62 actual cases within 90% CI and forecasts reflect the significant increase in the trend  
63 of global cases with growing uncertainty [43]. Forecasting and nowcasting domestic  
64 and international COVID-19 cases is also done with epidemiological compartmental  
65 models, like Susceptible-Exposed-Infectious-Recovered (SEIR) and SIR models  
66 [14,45,56]. Modern machine learning techniques like neural networks and black-box  
67 deep learning models were also used for COVID-19 forecasting [50,52]. However,  
68 forecasting COVID-19 cases is harder and this is primarily due to the following major  
69 factors [28]: (a) Very less amount of data is available; (b) Less understanding of the  
70 factors that contribute to it; (c) Model accuracy is constrained by our knowledge  
71 of the virus. With an emerging disease such as COVID-19, many biologic features  
72 of transmission are hard to measure and remain unknown; (d) Another source of  
73 uncertainty, affecting all models, is that we don't know how many people are, or have  
74 been, infected; (e) We are certainly missing a substantial number of cases due to vi-  
75 rologic testing, so models fitted to confirmed cases are likely to be highly uncertain [21].  
76

77 Time-series forecasting models take input of historical observations and extrapolate  
78 the patterns into the future. The past in no way will continue in the future for this  
79 pandemic, so it is challenging to model and produce predictions. There are essentially  
80 two general approaches to forecasting a time series: (a) generating forecasts from a  
81 single model; and (b) combining forecasts from many models (Hybrid). In classical  
82 time series forecasting, the ARIMA model is used predominantly for forecasting linear  
83 time series [4], which has a significant strong assumption of linearity in the system  
84 and homoscedastic error distribution; typically without sudden jumps and bursts.  
85 Individual models such as ARIMA, Wavelet ARIMA (WARIMA)[11], and Theta  
86 method [1,26] are inadequate to model such situations. Usage of nonlinear time series  
87 techniques in infectious disease modeling has successfully demonstrated the success  
88 of Artificial Neural Networks (ANN) and Autoregressive Neural Networks (ARNN)  
89 [6,32,34,58]. The idea of hybridizing forecasting models is not a new concept and  
90 their empirical applications resulted in superior forecasts to their counterparts. Most  
91 recently, some hybrid models combining linear and nonlinear time series models are  
92 proposed for COVID-19 [9,16] and performed exceptionally well for predicting these  
93 epidemics.

94  
95 Motivated by this, this study considers the time series data sets of coronavirus  
96 confirmed cases that show non-linearity, non-stationary and non-Gaussian patterns,  
97 thus making decisions based on a discrete model critical and unreliable. Another dif-  
98 ficulty in COVID-19 data from the modeling aspect is the unavailability of sufficient  
99 data points, which generates biased predictions and estimates, which can be well ma-  
100 neuvered by neural nets [38]. Most of the relevant studies focused on the outbreak's  
101 short-term and long-term forecasts of reported confirmed cases have a broad range of  
102 fluctuations, wide confidence intervals, poorly reported data and model specifics, and  
103 highly optimistic predictive performance [57]. Hybridization of two or more models is  
104 the most common solution [36] for optimizing forecasting performance, and efficiency  
105 with unknown complete data characteristics [30]. The importance of hybrid method-  
106 ology with a fusion of linear and non-linear forecasting models becomes evident in  
107 tackling such dynamic/non-linear time series and its inbuilt time-changing variance,  
108 complex autocorrelation structure [9,42]. With the growing need for advanced parsimonious hybrid forecasting methods accompanied by precise accuracy and accurate  
109 forecasts, the main objectives of this study are:

- 111 (1) To propose a simple and computationally efficient hybrid forecasting model  
112 which generates out-of-sample forecasts (two months) for five profoundly affected  
113 countries (United States of America (USA), Brazil, India, UK, and Canada) with  
114 higher accuracy as compared to state-of-the-art forecasting methods..
- 115 (2) To prove the model's stationarity and ergodicity properties from a statistical  
116 point of view.
- 117 (3) To discuss the merits and future challenges that needs attention while working  
118 with the proposal for epidemiological forecasting. We also recommend policy-  
119 making decisions, resource allocation based on these forecasts.

120 Therefore, this study proposes a novel hybrid Theta Autoregressive neural network  
121 model (TARNN) model combining Theta and ARNN models that can capture  
122 complex COVID-19 data structures. The proposed TARNN model has easy inter-  
123 pretability, produces robust predictions, and is adaptable to seasonality. An RShiny  
124 application is also built for TARNN can help practitioners reproducing and updating

125 the forecasts as further data becomes available.

126

127 The rest of the paper is organized as follows: Section 2 describes the formulation of  
128 the proposed hybrid TARNN model. The ergodicity and stationarity of the proposed  
129 hybrid model are discussed in Section 3. In Section 4, we discuss the country-wise  
130 COVID-19 confirmed case data sets and experimental results. The discussions about  
131 the results and practical implications are given in Section 5.

## 132 2. Methodology

133 We start by discussing the single forecasting models to be used in the hybridization,  
134 followed by the detailed formulation of the proposed hybrid Theta-ARNN model.

### 135 2.1. Theta Method

136 The ‘Theta method’ or ‘Theta model’ is a univariate time series forecasting technique  
137 that performs particularly well in M3 forecasting competition and of interest to fore-  
138 casters [1]. The method decomposes the original data into two or more lines, called  
139 theta lines, and extrapolates them using forecasting models. Finally, the predictions  
140 are combined to obtain the final forecasts. The theta lines can be estimated by simply  
141 modifying the ‘curvatures’ of the original time series. This change is obtained from a  
142 coefficient, called  $\theta$  coefficient, which is directly applied to the second differences of  
143 the time series:

$$Y_{new}''(\theta) = \theta Y_{data}'' \quad (1)$$

144 where  $Y_{data}'' = Y_t - 2Y_{t-1} + Y_{t-2}$  at time  $t$  for  $t = 3, 4, \dots, n$  and  $\{Y_1, Y_2, \dots, Y_n\}$  denote  
145 the observed univariate time series. In practice, the coefficient  $\theta$  can be considered as a  
146 transformation parameter which creates a series of the same mean and slope with that  
147 of the original data but having different variances. Now, Eqn. (1) is a second-order  
148 difference equation and has solutions of the following form [26]:

$$Y_{new}(\theta) = a_\theta + b_\theta(t - 1) + \theta Y_t, \quad (2)$$

149 where  $a_\theta$  and  $b_\theta$  are constants and  $t = 1, 2, \dots, n$ . Thus,  $Y_{new}(\theta)$  is equivalent to a  
150 linear function of  $Y_t$  with a linear trend added. The values of  $a_\theta$  and  $b_\theta$  are computed  
151 by minimizing the sum of squared differences:

$$\sum_{i=1}^t [Y_i - Y_{new}(\theta)]^2 = \sum_{i=1}^t [(1 - \theta)Y_i - a_\theta - b_\theta(i - 1)]^2. \quad (3)$$

152 Forecasts from the Theta model are obtained by a weighted average of forecasts of  
153  $Y_{new}(\theta)$  for different values of  $\theta$ . Furthermore, the prediction intervals and likelihood-  
154 based estimation of the parameters can be obtained based on a state space model  
155 which is demonstrated in [26]. The generalized version of the Theta method is suitable  
156 for automatic forecasting of time series [51].

## 157 **2.2. ARNN Model**

158 Artificial Neural Network-based forecasting methods received increasing interest in  
159 various applied domains in the late 1990s. A wide variety of neural nets are popularly  
160 used for supervised classification, prediction, and nonlinear time series forecasting [59].  
161 The architecture of a simple feedforward neural network can be described as a network  
162 of neurons arranged in an input layer, hidden layer, and output layer in a prescribed  
163 order. Each layer passes the information to the next layer using weights that are  
164 obtained using a learning algorithm [15]. ARNN model is a modification to the simple  
165 ANN model especially designed for prediction problems of time series data sets [15].  
166 ARNN model uses a prespecified number of lagged values of the time series as inputs  
167 and the number of hidden neurons in its architecture is also fixed [24]. ARNN( $p, k$ )  
168 model uses  $p$  lagged inputs of the time series data in a one hidden layered feedforward  
169 neural network with  $k$  hidden units in the hidden layer. Let  $\underline{x}$  denotes a  $p$ -lagged inputs  
170 and  $f$  is a neural network of the following architecture:

$$f(\underline{x}) = c_0 + \sum_{j=1}^k w_j \phi(a_j + b'_j \underline{x}); \quad (4)$$

171 where  $c_0, a_j, w_j$  are connecting weights,  $b_j$  are  $p$ -dimensional weight vector and  $\phi$  is  
172 a bounded nonlinear sigmoidal function (e.g., logistic squasher function or tangent  
173 hyperbolic activation function). These Weights are trained using a gradient descent  
174 backpropagation [49]. Standard ANN faces the dilemma to choose the number of hid-  
175 den neurons in the hidden layer and the optimal choice is unknown. However, for  
176 ARNN model, we adopt the formula  $k = [(p + 1)/2]$  for non-seasonal time series data  
177 where  $p$  is the number of lagged inputs in an autoregressive model [24].

## 178 **2.3. Proposed TARNN Model**

179 In this section, we describe the proposed hybrid model based on Theta method and  
180 ARNN model and we name it TARNN model. The proposed TARNN model is based  
181 on an error remodeling approach and there are broadly two types of error calculations  
182 popular in the literature which are given below [35].

**Definition 2.1.** In the additive error model, the forecaster treats the expert's estimate as a variable,  $\hat{Y}_t$ , and thinks of it as the sum of two terms:

$$\hat{Y}_t = Y_t + e_t,$$

183 where  $Y_t$  is the true value and  $e_t$  is the additive error term.

**Definition 2.2.** In the multiplicative error model, the forecaster treats the expert's estimate  $\hat{Y}_t$  as the product of two terms:

$$\hat{Y}_t = Y_t \times e_t,$$

184 where  $Y_t$  is the true value and  $e_t$  be the multiplicative error term.

185 Now, even if the relationship is of product type, in the log-log scale, it becomes  
186 additive. Hence, without loss of generality, we may assume the relationship to be

187 additive and expect the errors (additive) of a forecasting model to be random shocks.  
188 However, this is violated when there are complex correlation structures in the time  
189 series data and less amount of knowledge is available about the data generating process.  
190 A simple example is the daily confirmed cases of the COVID-19 cases for various  
191 countries where very little is known about the structural properties of the current  
192 pandemic. Thus, we need two-stage modeling approach to formulate this complex  
193 time series problem. The proposed TARNN model is a hybrid model based on the  
194 additive error re-modeling approach. The hybrid TARNN approach consists of three  
195 basic steps:

- 196 • In the first step of the TARNN model, the Theta method is applied to the time  
197 series data to model the linear components of a given time series data set.
- 198 • Theta model generates in-sample forecasts and the error series is calculated.
- 199 • In the next phase, the residuals (additive errors) generated by the Theta method  
200 are remodeled using a nonlinear ARNN model. Finally, both the forecasts obtained  
201 from the Theta and ARNN models are combined together to get the final  
202 forecasts for the given time series.

203 The mathematical formulation of the proposed hybrid TARNN model ( $Z_t$ ) is as  
204 follows:

$$Z_t = L_t + N_t,$$

205 where  $L_t$  is the linear part and  $N_t$  is the nonlinear part of the hybrid model. We can  
206 estimate both  $L_t$  and  $N_t$  from the available time series data. Let  $\hat{L}_t$  be the forecast  
207 value of the Theta model at time  $t$  and  $\epsilon_t$  represent the error residuals at time  $t$ ,  
208 obtained from the Theta model. Then, we write

$$\epsilon_t = Z_t - \hat{L}_t.$$

209 These left-out residuals are further modeled by ARNN model and can be represented  
210 as follows:

$$\epsilon_t = f(\epsilon_{t-1}, \epsilon_{t-2}, \dots, \epsilon_{t-p}) + \varepsilon_t,$$

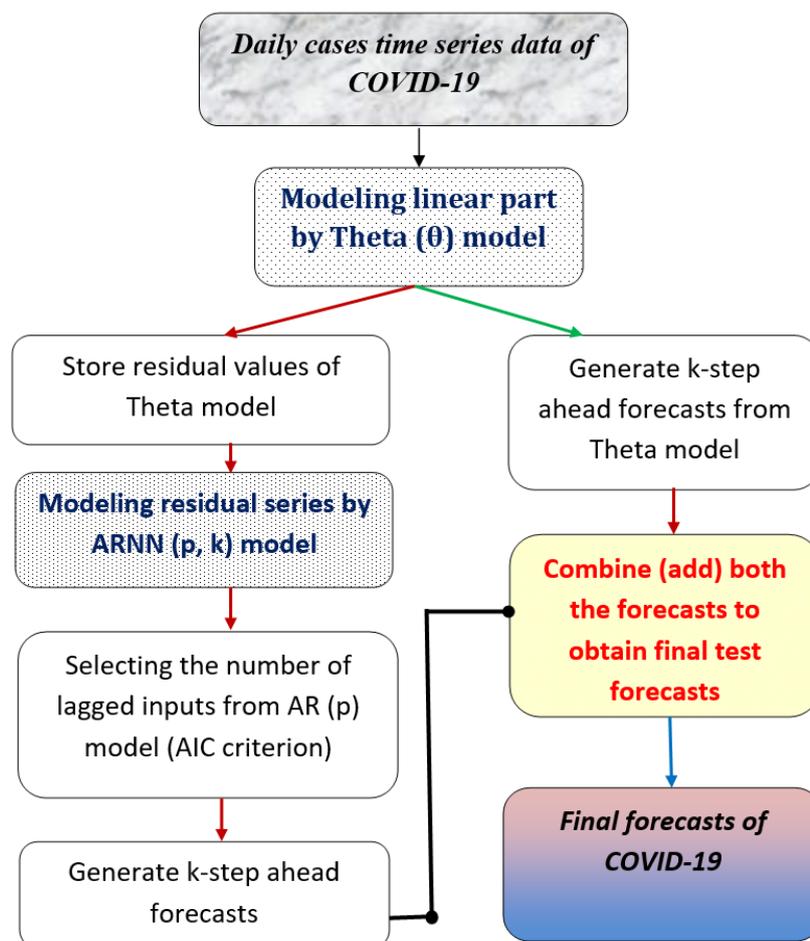
211 where  $f$  is a nonlinear function and the modeling is done by the ARNN model as  
212 defined in Eqn. (4) and  $\varepsilon_t$  is supposed to be random shocks. Therefore, the combined  
213 forecast can be obtained as follows:

$$\hat{Z}_t = \hat{L}_t + \hat{N}_t,$$

214 where  $\hat{N}_t$  is the forecasted value of the ARNN model. An overall flow diagram of the  
215 proposed TARNN model is given in Figure 1.

216

217 In the proposed TARNN model, ARNN is applied to remodel the left-over autocor-  
218 relations in the residuals which Theta model individually could not model. Thus, the  
219 proposed TARNN model can be considered as an error re-modeling approach. This is  
220 important because due to model mis-specification and disturbances in the pandemic  
221 time series, the linear Theta model may fail to generate white noise behavior for the  
222 forecast residuals. The TARNN approach eventually can improve the predictions for  
223 the epidemiological forecasting problem as shown in Section 4.



**Figure 1.** The main workflow of the proposed TARNN model.

224 **Remark 1.** The idea of additive error modelling is useful for modeling complex time  
 225 series for which achieving random shocks based on individual forecasting models is  
 226 difficult. More precisely, the TARNN approach is developed for forecasting the COVID-  
 227 19 confirmed cases for which the data generating process and the various characteristics  
 228 of the epidemic are still unknown. The proposed TARNN model only assumes that  
 229 the linear and nonlinear components of the epidemic time series can be separated  
 230 individually.

### 231 3. Ergodicity and Stationarity of the proposed TARNN model

232 In this section, we derive the results for the ergodicity and asymptotic stationarity  
 233 of the proposed TARNN model. The ergodicity and stationarity is of particular  
 234 importance from a statistician's point of view in time series analysis since for such  
 235 processes a single realization displays the whole probability of the data generating  
 236 process. We use several previous results on nonlinear time series and Markov chains  
 237 to find sufficient conditions for which the overall process is ergodic and stationary  
 238 [5,7,10,53].

239

To start, we write the underlying stochastic model of the Theta method using the state space approach. We initialize the model by setting  $Y_1 = l_1$  and then for  $t = 2, 3, \dots$ ; and drift term  $b$ , let

$$Y_t = l_{t-1} + b + \epsilon_t \quad \text{and} \quad l_t = l_{t-1} + b + \alpha \epsilon_t ;$$

240 where  $\{\epsilon_t\}$  is assumed to follow Gaussian white noise with mean zero and variance  $\sigma^2$   
 241 and  $\alpha$  is the smoothing parameter for the simple exponential smoothing (SES) model.  
 242 Now,  $Y_t$  follows a state space model which gives forecasts equivalent to SES with drift  
 243 [26]. Also,  $Y_t$  can be written in the following form:

$$Y_t = Y_{t-1} + b + (\alpha - 1)\epsilon_{t-1} + \epsilon_t. \quad (5)$$

244 The above is an ARIMA(0,1,1) process with drift term [26]. The left-out residuals of  
 245 Theta model is further modeled by ARNN process. We consider the ARNN process  
 246 generated by the additive noise of the ARIMA(0,1,1) process with drift. Let  $\epsilon_t$  be the  
 247 process defined by the stochastic difference equation of the following form:

$$\epsilon_t = f(\epsilon_{t-1}, \epsilon_{t-2}, \dots, \epsilon_{t-p}, \Theta) + \varepsilon_t, \quad (6)$$

248 where  $\varepsilon_t$  is an i.i.d. noise process and  $f(\cdot, \Theta)$  is a feedforward neural network with  
 249 weight (parameter) vector  $\Theta$  and inputs  $\epsilon_{t-1}, \epsilon_{t-2}, \dots, \epsilon_{t-p}$ . The definition of  $f$  is  
 250 given in Eqn. (4).

### 251 **3.1. Time Series as Markov chains**

252 We start by defining the following notations:

$$\begin{aligned} z_{t-1} &= (\epsilon_{t-1}, \dots, \epsilon_{t-p})' \\ F(z_{t-1}) &= (f(z_{t-1}), \epsilon_{t-1}, \dots, \epsilon_{t-p+1})' \\ e_t &= (\varepsilon_t, 0, \dots, 0)' \end{aligned}$$

253 Then Eqn. (6) can be written as follows [10]:

$$z_t = F(z_{t-1}) + e_t \quad (7)$$

254 with  $z_t, e_t \in \mathbb{R}^p$ . In this section, we show the (strict) stationarity of the state space  
 255 form, as defined in Eqn. (7). The problem of showing  $\{z_t\}$  to be stationary is closely  
 256 related to the ergodicity of the process [55]. A Markov chain  $\{z_t\}$  is called geometrically  
 257 ergodic if there exists a probability measure  $\pi$  on the state space  $(\mathbb{R}^p, \mathbb{B})$  and a constant  
 258  $\rho > 1$  such that

$$\lim_{n \rightarrow \infty} \rho^n \|P^n(z, A) - \pi(A)\| = 0 \quad (8)$$

259 for each  $z \in \mathbb{R}^p$ ,  $\mathbb{B}$  is the Borel  $\sigma$ -algebra on  $\mathbb{R}^p$  and  $\|\cdot\|$  denotes the total variation  
 260 norm. If Eqn. (8) holds for  $\rho = 1$ , then  $\{z_t\}$  is called ergodic.

261

The definition for  $P^n(z, A)$  can be given as the probability that  $\{z_n\}$  moves from  $z$  to the set  $A \in \mathbb{B}$  in  $n$  steps:  $P^n(z, A) = P(z_{t+n} \in A | z_t = z)$ . Also, expression for  $\pi(A)$  is as follows:

$$\pi(A) = \int P^1(z, A)\pi(dz) \text{ for all } A \in \mathbb{B}.$$

Thus, we call  $\pi$  as the stationary measure and the distribution of  $z_t$  converges to  $\pi$  if  $\{z_t\}$  is ergodic. Then, we say  $\{z_t\}$  is asymptotically stationary [33]. To establish the ergodicity of TARNN processes, we need the concept of irreducibility and aperiodicity. A Markov chain  $\{z_t\}$  is called irreducible if

$$\sum_{n=1}^{\infty} P^n(z, A) > 0 \text{ for all } z \in \mathbb{R}^p,$$

whenever  $\lambda(A) > 0$  and  $\lambda$  denotes the Lebesgue measure on  $(\mathbb{R}^p, \mathbb{B})$ . Thus, for an irreducible Markov chain, all parts of the state space can be reached by the Markov chain irrespective of the starting point. Now, an irreducible Markov chain is aperiodic if there exists an  $A \in \mathbb{B}$  with  $\lambda(A) > 0$  and for all  $C \in \mathbb{B}$ ,  $C \subseteq A$  with  $\lambda(C) > 0$ , there exists a positive integer  $n$  such that

$$P^n(z, C) > 0 \text{ and } P^{n+1}(z, C) > 0 \text{ for all } z \in \mathbb{R}^p.$$

262 Hence, it is possible that the Markov chain returns to given sets only at specific time  
 263 points for an aperiodic Markov chain. For most general time series models, irreducibil-  
 264 ity and aperiodicity cannot be assumed automatically. However, for a TARNN process,  
 265 these conditions can be checked. In general, it is sufficient to assume the distribution of  
 266 the noise process to be an absolutely continuous component with respect to Lebesgue  
 267 measure and the support of the probability density function (PDF) is sufficiently large.

### 268 3.2. Main Results

269 It is clear from the above discussion that if the Markov chain is geometrically ergodic,  
 270 then its distribution will converge to  $\pi$  and the corresponding time series will be called  
 271 asymptotically stationary, see also [53]. Lemma 3.1 states that the state space of the  
 272 Markov chain cannot be reduced depending on the starting point.

273 **Lemma 3.1.** *Let  $\{z_t\}$  is defined by (7), and let  $E|\varepsilon_t| < \infty$  and the PDF of  $\varepsilon_t$  is positive  
 274 everywhere in  $\mathbb{R}$ . Then if  $f$  is defined by (4), the Markov chain  $\{z_t\}$  is  $\varphi$ -irreducible  
 275 and aperiodic.*

276 **Proof.** Since the support of the PDF of  $\varepsilon_t$  is the whole real line, that is, the PDF  
 277 is positive everywhere in  $\mathbb{R}$ , then we can say that  $\{z_t\}$  is  $\varphi$ -irreducible by using [10].  
 278 In our case, every non-null  $p$ -dimensional hypercube can be reached in  $p$  steps with  
 279 positive probability (and hence every non-null Borel set  $A$ ). A necessary and sufficient  
 280 condition for  $\{z_t\}$  to be aperiodic is to have a set  $A$  and positive integer  $n$  such that  
 281  $P^n(z, A) > 0$  and  $P^{n+1}(z, A) > 0$  for all  $z \in A$  [53]. In this case, this is true for all  $n$   
 282 due to consideration of the unbounded additive noise.  $\square$

283 The theorem below states the necessary conditions for geometric ergodicity of a  
 284 Markov chain. This can be obtained using the decomposition technique and ergodicity  
 285 of stochastic difference equations [7,10].

286 **Theorem 3.2.** *Suppose  $\{z_t\}$  is defined as in (6) and (7),  $F$  be a compact set that can  
 287 be decomposed as  $F = F_h + F_d$ , and the following conditions hold:*

- 288 (i)  $F_h(\cdot)$  is continuous and homogeneous and  $F_d(\cdot)$  is bounded;  
 289 (ii)  $E|\varepsilon_t| < \infty$  and probability distribution function of  $\varepsilon_t$  is positive everywhere in  $\mathbb{R}$ ;  
 290 then  $\{z_t\}$  is geometrically ergodic.

291 **Proof.**  $\{z_t\}$  satisfies the following equation:

$$z_t = F(z_{t-1}) + e_t; \quad t \geq 1, \text{ and } F: \mathbb{R}^p \rightarrow \mathbb{R}^p. \quad (9)$$

292 Let  $F_h$  be continuous and homogeneous, viz.,  $F_h(cz) = cF_h(z)$  for all  $c > 0$ ,  $z_t \in \mathbb{R}^p$ ,  
 293 and  $F_d$  is bounded. Let the origin,  $O$ , be a fixed point of  $F_h$ . It is important to note that  
 294  $\varepsilon_t$  satisfies the condition (ii) in Theorem 3.2. We are going to show that the existence  
 295 of a continuous Lyapunov function,  $V$ , in a neighbourhood of the origin which will  
 296 ensure the geometric ergodicity of (7).

297 To start with we let  $W \subseteq \mathbb{R}^p$ , the closure of  $W$  by  $\overline{W}$  and its boundary by  $\widehat{W}$ . We also  
 298 let  $V$  be defined over the closure of the unit ball. We let  $p_0 = \inf_{\|z\|=1} V(z)$ , where  
 299  $\|\cdot\|$  denote the Euclidean norm. Also, let  $G$  be the maximal connected component of  
 300  $\{z: V(z) < \frac{p_0}{2}\}$  that contains the origin. Then we have  $F(\overline{G}) \subset G$ .

301 Let  $g(z) = \inf\{r \geq 0, z \in rG\}$ ,  $z \in \mathbb{R}^p$ , where  $rG = \{rz, z \in G\}$ . Then  $g(z)$  is well  
 302 defined and it can easily be checked that  $g$  has the following properties:

- 303 (1)  $g(cz) = cg(z)$ , for all  $c > 0$ .  
 304 (2) There exists  $0 < c < \mathbb{C} < \infty$  such that  $c\|z\| \leq g(z) \leq \mathbb{C}\|z\|$ .  
 305 (3)  $(z/g(z)) \in \widehat{G}$ .  
 306 (4) There exists  $\epsilon > 0$ ,  $0 < \theta < 1$  such that for all  $z \in \widehat{G}$ ,  $y \in \mathbb{R}^p$ , we have  
 307  $\|y - F(z)\| < \epsilon \Rightarrow y \in G$  and  $g(y) < \theta$ .

Now, for Eqn. (7),  $\varepsilon_t$  satisfies  $E|\varepsilon_t| < \infty$ . Let  $A \in \mathbb{B}$  and  $z \in \mathbb{R}^p$ , we define  $P(z, A)$  be the transition probability function as:

$$P(z, A) = \int_{A-F(z)} f(t)\mathbb{P}(dt).$$

Thus, it holds that

$$\begin{aligned}
 & \int g(y)P(z, dy) \\
 &= \int g(F_h(z) + t)f(t)\mathbb{P}(dt) \\
 &= \int_{\|t/g(z)\| < \epsilon} g(z)g\left(F_h\left(\frac{z}{g(z)}\right) + \frac{t}{g(z)}\right)f(t)\mathbb{P}(dt) \\
 & \quad + \int_{\|t/g(z)\| \geq \epsilon} g(z)g\left(F_h\left(\frac{z}{g(z)}\right) + \frac{t}{g(z)}\right)f(t)\mathbb{P}(dt) \\
 &< [\theta g(z)] + [M + g(z)\beta(z)] \text{ for some } M > 0.
 \end{aligned}$$

Here  $|\beta(z)| < B < \infty$  for all  $z$  and  $\beta(z) \rightarrow 0$  as  $\|z\| \rightarrow \infty$ . We let  $h(z) = g(z) + 1$ , and  $r$  be such that

$$|\beta(z)| < \frac{(1 - \theta)}{4} \quad \text{for } \|z\| > r.$$

308 Then for  $r_0 = \max\left(r, \frac{4(C+1)}{(1-\theta)^p}\right)$ , there exists  $B'$  such that

- 309 (1)  $\int f(y)P(z, dy) < B' < \infty$  when  $\|z\| \leq r_0$ .  
 310 (2)  $\int f(y)P(z, dy) < \frac{1}{2}(1 + \theta)h(z)$  when  $\|z\| > r_0$ .

311 Using Theorem 4 of [55], we can conclude that  $\{z_t\}$  is geometric ergodic.  $\square$

312 The next theorem gives the main result for asymptotic stationary of the TARNN  
 313 model.

314 **Theorem 3.3.** *Let  $E|\varepsilon_t| < \infty$  and the PDF of  $\varepsilon_t$  is positive everywhere in  $\mathbb{R}$ , and  $\{\varepsilon_t\}$   
 315 and  $\{z_t\}$  are defined as in (6) and (7), respectively. Then if  $f$  is a nonlinear neural  
 316 network as defined in (4), then  $\{z_t\}$  is geometrically ergodic and  $\{\varepsilon_t\}$  is asymptotically  
 317 stationary.*

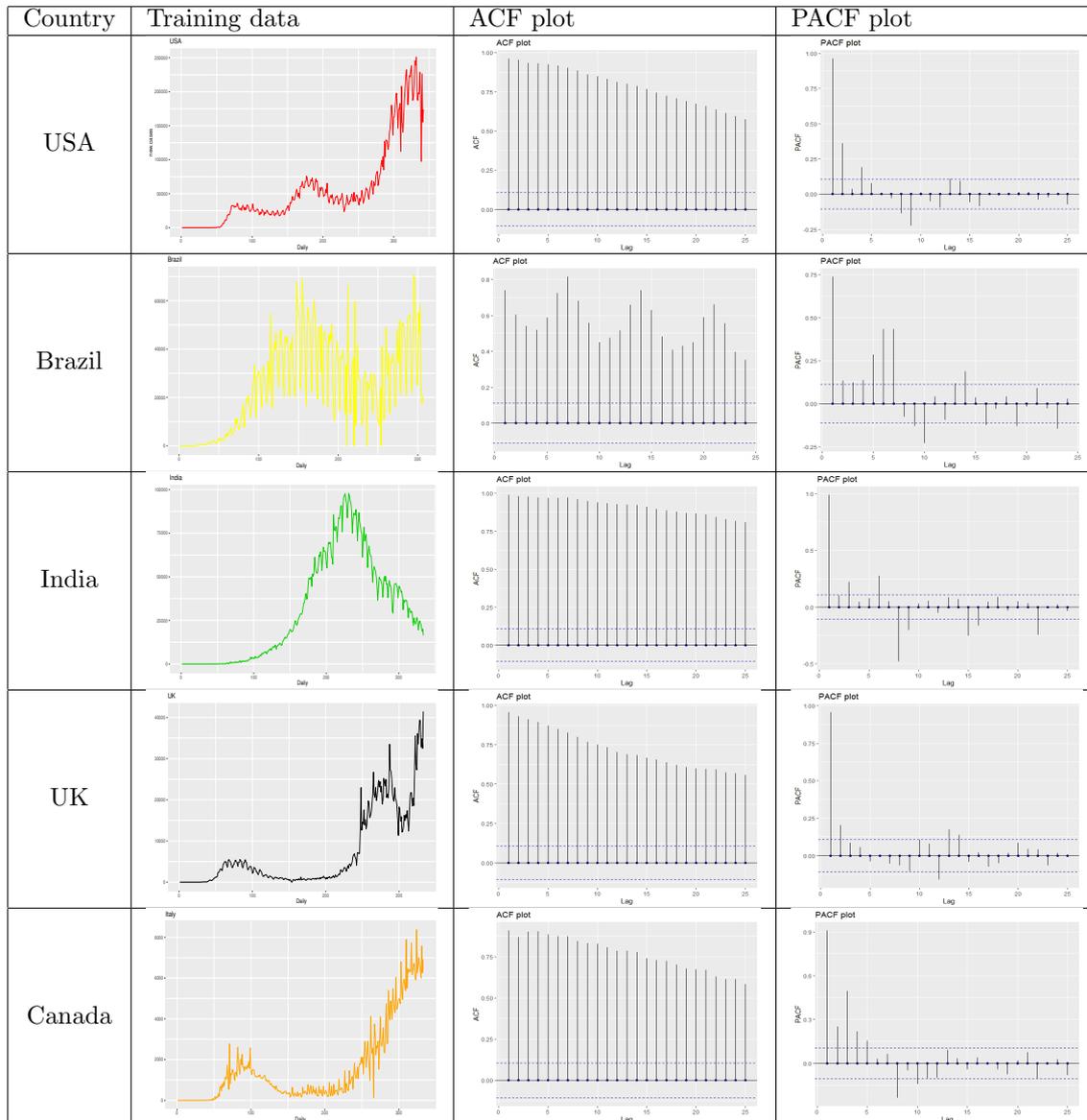
318 **Proof.** The noise process  $\varepsilon_t$  satisfies  $E|\varepsilon_t| < \infty$  by assumption (e.g., Gaussian noise).  
 319 It is also important to note that neural network activation functions, more precisely  
 320 logistics or tan-hyperbolic activation functions, are continuous compact functions and  
 321 have a bounded range. Thus  $\{z_t\}$  satisfies all the criteria to be geometrically ergodic  
 322 and using Theorem (3.2), one can write that for the ARNN process with  $F_h \equiv 0$  and  
 323  $F_d \equiv F$ . Thus, the series  $\{\varepsilon_t\}$  is asymptotically stationary.  $\square$

324 **Remark 2.** Some interpretations and practical implications of the theoretical results  
 325 are given below:

- 326 • The geometric rate of convergence in Theorem 3.3 implies that the memory  
 327 of TARNN process vanishes exponentially fast. This implies that the simplest  
 328 version of the proposed model converges to a Wiener process.
- 329 • This is important for predictions over larger intervals of time, for example, one  
 330 might train the network on an available sample and then use the trained network  
 331 to generate new data with similar properties like the training sample. The results

332 for the asymptotic stationarity can guarantee that the proposed hybrid model  
 333 can not have growing variance over time.  
 334 • From practitioners point of view, when the data is generated by the irreducible  
 335 TARNN process, the estimated weights are not too far from the true weights.  
 336 Then, one can draw an indirect conclusion on the statistical nature of the ob-  
 337 served data based on the estimated weights.

#### 338 4. Experimental Analysis



**Table 1.** Training data sets (new daily cases) and corresponding ACF, PACF plots for all the countries considered in this study.

339 Five time series COVID-19 datasets are considered for assessing various forecasting  
 340 models (individual and hybrid). The datasets are mostly nonlinear, nonstationary, and

341 non-Gaussian in nature. We have used root mean square error (RMSE), mean absolute  
342 error (MAE), and mean absolute scaled error (MASE) [27] to evaluate the predictive  
343 performance of the models used in this study. Since the number of data points in all  
344 the datasets is limited, advanced deep learning techniques will overfit the datasets [20].

#### 345 **4.1. Data Description**

346 Data is collected from the “Our World in Data” public repository (Link: <https://ourworldindata.org/coronavirus>) on five countries: USA, Brazil, India, UK, and  
347 Canada that have the leading number of confirmed cases of COVID-19. These are  
348 univariate time-series from inception-Feb 26, 2021 of confirmed cases that are analyzed  
349 for generating future outbreak predictions. We have taken last 60 data points as test  
350 points on which the forecasting models are evaluated. A summary of the COVID-19  
351 data sets of confirmed cases and their descriptions are presented in Table 2. A pictorial  
352 view of the training data sets along with auto-correlation function (ACF) and partial  
353 ACF (PACF) plots are given in Table 1.

Country Name	Start-End Date	No. of Days	Minimum Value	Maximum Value	Train size	Test size
USA	22/01/2020 - 26/02/2021	402	0	251161	342	60
Brazil	26/02/2020 - 26/02/2021	367	0	70574	307	60
India	30/01/2020 - 26/02/2021	394	0	97894	334	60
UK	31/01/2020 - 26/02/2021	393	0	41460	333	60
Canada	26/01/2020 - 26/02/2021	398	0	10100	338	60

**Table 2.** Description of COVID-19 datasets.

#### 355 **4.2. Performance Evaluation Metrics**

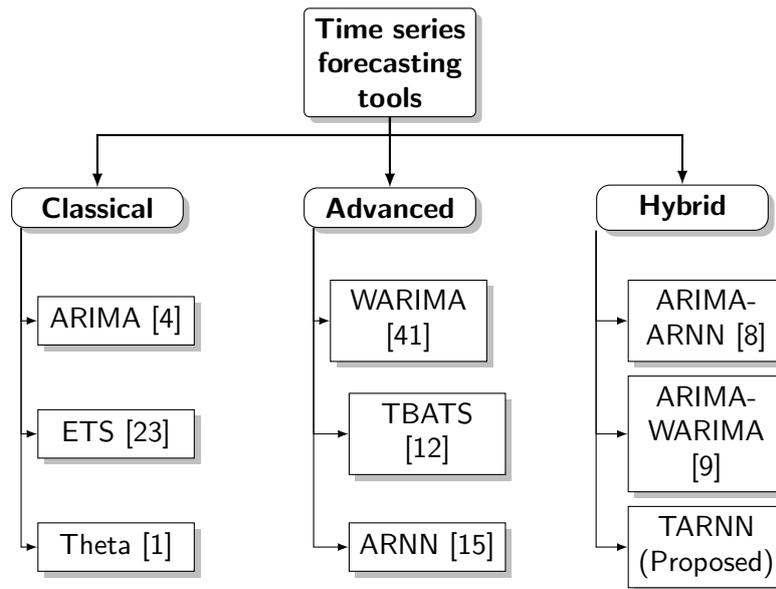
The performance of different forecasting models are evaluated based on RMSE, MAE, and MASE metrics for these five time series [24]:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}; \quad MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

356 where,  $y_i$  is the actual value,  $\hat{y}_i$  is the predicted value, and  $n$  denotes the number of  
357 data points. MASE is computed using ‘mase’ function available under the “Metrics”  
358 package in R statistical software. By definition, the lower the value of these perfor-  
359 mance metrics, the better is the performance of the concerned forecasting model.

#### 360 **4.3. Analysis of Results**

361 A schematic diagram is presented in Figure 2 to give an outline of the models  
362 to be used in this section. We start the experimental evaluation for all data sets  
363 with the classical ARIMA( $p, d, q$ ) using ‘forecast’ [25] statistical package in the R  
364 statistical software [44]. The nonlinear and nonseasonal time series of these countries  
365 were modelled with traditional and advanced methods such as ARIMA, Theta,



**Figure 2.** Time series forecasting tools (available and proposed) used in this study.

366 WARIMA, ETS, TBATS, ARNN, and two popularly used hybrid models, namely  
 367 hybrid ARIMA-ARNN [8], hybrid ARIMA-WARIMA [9], and the newly introduced  
 368 hybrid Theta-ARNN (TARNN) models. For the accuracy of prediction, the data was  
 369 partitioned into training and test sets, where the test set consisted of last 60 data  
 370 points, and the training set included (inception - 60+1) data points. Table 3 gives the  
 371 essential details about the functions and packages required for the implementation of  
 372 standard individual models.

373

374 In the proposed TARNN model, linear modelling is conducted with Theta model  
 375 using ‘thetaf’ function under the “forecast” package in R statistical software.  
 376 Nonlinear modelling with ARNN approach is carried out with “caret” package using  
 377 ‘nnetar’ function in R statistical software. Furthermore, Theta residuals are modelled  
 378 with  $ARNN(p, k)$  model having a pre-defined Box-Cox transformation set  $\lambda = 0$  to  
 379 ensure the forecast values to stay positive. The values of  $p$  and  $k$  are obtained by  
 380 training the network, which is a data-dependent approach as discussed in Section  
 381 2.2. Further, both the linear and nonlinear forecasts are summed up to obtain  
 382 the final forecasts. Theta model was fitted to five data sets namely USA, Brazil,  
 383 India, UK, and Canada. Theta model residuals for these five countries were trained  
 384 using  $ARNN(15,8)$ ,  $ARNN(21,11)$ ,  $ARNN(25,13)$ ,  $ARNN(24,12)$ ,  $ARNN(12,6)$ , and  
 385  $ARNN(9,5)$  models with an average of 500 networks for all five data sets, each of  
 386 which is a 15-8-1, 21-11-1, 25-13-1, 24-12-1, 12-6-1, and 9-5-1 networks with 137,  
 387 254, 352, 313, 85, and 56 weights and with 5807711, 309816, 20735, 883.6, 209396,  
 388 and 59477 estimated  $\sigma^2$ , respectively. Finally, the predicted results of both Theta  
 389 and ARNN models are added together to obtain the estimated forecasts of the pro-  
 390 posed TARNN model. Rest of the hybrid models were compared in the similar fashion.

391

392 We compared our proposed TARNN model with traditional single models  
 393 (ARIMA, TBATS, Theta, ARNN, WARIMA, ETS) along with hybrid ARIMA-  
 394 WARIMA model, hybrid ARIMA-ARNN model and the experimental results are

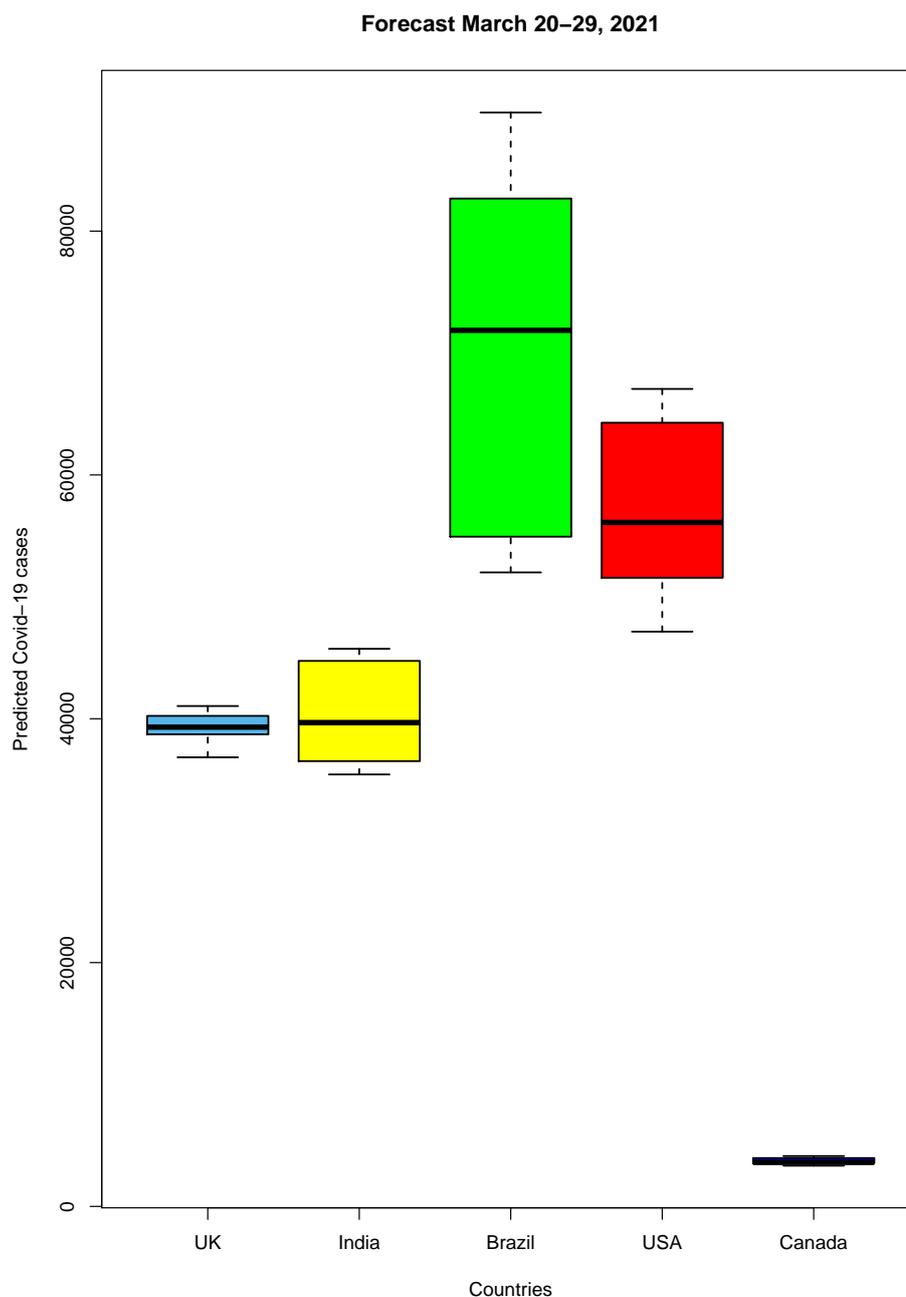
Model	R function	R package	Reference
ARIMA	auto.arima	forecast	[25]
ETS	ets	forecast	[25]
TBATS	tbats	forecast	[25]
Theta	thetaf	forecast	[25]
ARNN	nnetar	forecast	[25]
WARIMA	WaveletFittingarma	WaveletArima	[39]

**Table 3.** R functions and packages for standard forecasting model implementations.

395 reported in Table 4. The performance of the proposed hybrid Theta-ARNN (TARNN)  
 396 model is superior compared to all traditional individual and hybrid models on average.  
 397 The theoretically proven asymptotic stationarity of the proposed hybrid model also  
 398 suggests that the model cannot have a growing variance over time. The consistency  
 399 and adequacy in the experimental results empirically approves the same. Thus, the  
 400 efficacy of the proposed methodology of the proposed hybrid model is experimentally  
 401 validated. Nevertheless, no model can have dominant advantage and this is related  
 402 to no free lunch theorem in statistical learning [20]. Forecast of COVID-19 cases for  
 403 March 20-29, 2021 are generated using proposed TARNN model and shown in Figure 3.  
 404

Methods	USA	UK	India	Canada	Brazil
MASE values					
ARIMA	3.55	5.53	0.67	2.08	1.55
Theta	3.24	5.88	1.89	2.08	1.31
WARIMA	4.97	6.46	5.26	1.8	1.53
ETS	2.47	5.62	1.17	1.65	1.86
TBATS	<b>2.46</b>	5.58	1.44	1.81	4.77
ARNN	3.29	5.67	3.16	2.23	2.83
ARIMA-ARNN	3.59	5.50	<b>0.71</b>	2.06	1.53
ARIMA-WARIMA	3.64	5.69	0.72	2.05	1.58
Proposed TARNN	3.27	<b>5.07</b>	1.60	<b>1.53</b>	<b>1.05</b>
RMSE values					
ARIMA	85378.21	19598.5	5325.44	2879.54	28661.69
Theta	77538.27	21033.15	9159.85	2976.44	22950.91
WARIMA	116791.17	43298.22	25609.64	2764.52	28139.06
ETS	60473.05	20017.92	6527.45	2425.78	33242.02
TBATS	<b>60472.43</b>	19855.08	7262.95	2647.24	86202.46
ARNN	79694.81	16796.84	14859.34	3063.08	48693.24
ARIMA-ARNN	86626.36	19557.8	<b>5323.38</b>	2876.86	27825.87
ARIMA-WARIMA	86938.13	20249.21	5428.55	2845.54	28908.41
Proposed TARNN	78320.37	<b>16559.04</b>	7854.66	<b>2323.95</b>	<b>20423.69</b>
MAE values					
ARIMA	72823.9	17672.58	2833.76	2390.47	24714.92
Theta	66491.15	18797.76	8024.95	2497.19	20853.47
WARIMA	101923.84	36890.4	22279.76	2069.01	24294.93
ETS	50537.67	17953.84	4972.05	1893.77	29506.42
TBATS	<b>50533.39</b>	17839.37	6091.65	2074.26	75834.12
ARNN	67409.69	14941.82	13389.13	2556.93	44939.99
ARIMA-ARNN	73690.5	17585.56	<b>3023.91</b>	2386.66	24259.25
ARIMA-WARIMA	74637.97	18187.05	3056.05	2353.5	25162.57
Proposed TARNN	66939.65	<b>14770.08</b>	5487.11	<b>758.82</b>	<b>16701.18</b>

**Table 4.** Quantitative measure of performances of forecasting methods on the COVID-19 test data sets for five countries. Best results are made 'bold'. Proposed TARNN beats all the models for three out of five data sets and remain competitive for other two data sets.



**Figure 3.** Forecast of COVID-19 cases for March, 2021 using proposed TARNN model.

## 405 5. Discussions and Conclusions

406 In this study, we proposed a novel hybrid Theta-ARNN (TARNN) model using  
407 residual modelling approach that performs considerably well for confirmed cases of  
408 COVID-19 forecasting for the countries that include the ones with the highest number  
409 of cases USA, followed by Brazil, India, UK, and Canada. The proposed TARNN

410 model filters linearity using the Theta model and can better explain the linear,  
411 nonlinear, and nonstationary tendencies present in the selected COVID-19 data sets  
412 compared to the traditional single and hybrid models. It also yields better forecast  
413 accuracy than various traditional single and hybrid models for three out of five  
414 countries. The proposal will be useful in decision and policy making for government  
415 officials and policymakers to allocate adequate health care resources for the coming  
416 days in responding to the crisis. Time series of epidemics can oscillate heavily due to  
417 various epidemiological factors and these fluctuations are challenging to be captured  
418 adequately for precise forecasting. This newly developed model can still predict with  
419 better accuracy provided the conditions of asymptotic stationarity of the hybrid  
420 model are satisfied. This method can be used to update real-time forecasts as more  
421 data becomes available. The study covering multiple countries can be utilized without  
422 geographical borders and reflects the impact of social distancing, wearing masks,  
423 lock down, shutdown, quarantine, and sanitizing proper measures implemented by  
424 authorities. Prevalent techniques in literature were unable to completely capture  
425 the nonlinear behavior of stochastic time series containing inherent random shock  
426 components. This new method have significant theoretical (established ergodicity  
427 and stationarity of the proposed TARNN process) as well as practical implications.  
428 Authorities and health care can modify their planning in stockpiles and hospital-beds  
429 depending on these forecasts of the COVID-19 pandemic.

430

431 Many parameters associated with COVID-19 transmission are still poorly under-  
432 stood. The resulting model uncertainty is not always calculated or reported in a stan-  
433 dardized way. Once we can incorporate these variables, we can improve our estimates  
434 and update the TARNN model accordingly. Since purely statistical approaches do not  
435 account for how transmission occurs, they are generally not well suited for long-term  
436 predictions about epidemiological dynamics (such as when the peak will occur and  
437 whether a resurgence will happen) or for inferences about intervention efficacy. Most  
438 forecasting models therefore limit their projections to one week or a few weeks ahead.  
439 Moreover, the problem of using confirmed cases to fit models is further complicated  
440 by the fact that the fraction of cases that are confirmed is spatially heterogeneous and  
441 time-varying. Amid enormous uncertainty about the future of the COVID-19 pan-  
442 demic, the proposed TARNN model yields quantitative projections that policymakers  
443 may need in the short term to allocate resources or plan interventions. To conclude,  
444 this model can further be easily extended for similar nonlinear and non-Gaussian fore-  
445 casting problems arising in other applied domains.

## 446 **Data and Code**

447 All the results can be effortlessly updated with the help of R-shiny application as new  
448 data becomes available. This publicly available repository link [https://github.com/arinjita9/COVID-19-Forecasting-by-TARNN-](https://github.com/arinjita9/COVID-19-Forecasting-by-TARNN) contains the current data files and R  
449 scripts for the TARNN model which ensures the repeatability and reproducibility of  
450 the results presented in this study.  
451

## 452 **Disclosure statement**

453 The authors declare that they have no competing interests.

## 454 Funding

455 The authors had no funding for this work.

## 456 References

- 457 [1] V. Assimakopoulos and K. Nikolopoulos, *The theta model: a decomposition approach to*  
458 *forecasting*, International journal of forecasting 16 (2000), pp. 521–530.
- 459 [2] D. Baud, X. Qi, K. Nielsen-Saines, D. Musso, L. Pomar, and G. Favre, *Real estimates of*  
460 *mortality following covid-19 infection*, The Lancet infectious diseases (2020).
- 461 [3] M.N.K. Boulos and E.M. Geraghty, *Geographical tracking and mapping of coronavirus*  
462 *disease covid-19/severe acute respiratory syndrome coronavirus 2 (sars-cov-2) epidemic*  
463 *and associated events around the world: how 21st century gis technologies are supporting*  
464 *the global fight against outbreaks and epidemics* (2020).
- 465 [4] G.E. Box, G.M. Jenkins, G.C. Reinsel, and G.M. Ljung, *Time series analysis: forecasting*  
466 *and control*, John Wiley & Sons, 2015.
- 467 [5] P.J. Brockwell and A. Lindner, *Strictly stationary solutions of autoregressive moving av-*  
468 *erage equations*, Biometrika 97 (2010), pp. 765–772.
- 469 [6] E. Cadenas and W. Rivera, *Wind speed forecasting in three different regions of mexico,*  
470 *using a hybrid arima-ann model*, Renewable Energy 35 (2010), pp. 2732–2738.
- 471 [7] T. Chakraborty, A.K. Chakraborty, M. Biswas, S. Banerjee, and S. Bhattacharya, *Unem-*  
472 *ployment rate forecasting: A hybrid approach*, Computational Economics 57 (2021), pp.  
473 183–201.
- 474 [8] T. Chakraborty, S. Chattopadhyay, and I. Ghosh, *Forecasting dengue epidemics using*  
475 *a hybrid methodology*, Physica A: Statistical Mechanics and its Applications (2019), p.  
476 121266.
- 477 [9] T. Chakraborty and I. Ghosh, *Real-time forecasts and risk assessment of novel coronavirus*  
478 *(covid-19) cases: A data-driven analysis*, Chaos, Solitons & Fractals (2020), p. 109850.
- 479 [10] K.S. Chan and H. Tong, *On the use of the deterministic lyapunov function for the er-*  
480 *godicity of stochastic difference equations*, Advances in applied probability 17 (1985), pp.  
481 666–678.
- 482 [11] A.J. Conejo, M.A. Plazas, R. Espinola, and A.B. Molina, *Day-ahead electricity price*  
483 *forecasting using the wavelet transform and arima models*, IEEE transactions on power  
484 systems 20 (2005), pp. 1035–1042.
- 485 [12] A.M. De Livera, R.J. Hyndman, and R.D. Snyder, *Forecasting time series with complex*  
486 *seasonal patterns using exponential smoothing*, Journal of the American statistical asso-  
487 ciation 106 (2011), pp. 1513–1527.
- 488 [13] E.J. Emanuel, G. Persad, R. Upshur, B. Thome, M. Parker, A. Glickman, C. Zhang, C.  
489 Boyle, M. Smith, and J.P. Phillips, *Fair allocation of scarce medical resources in the time*  
490 *of covid-19* (2020).
- 491 [14] D. Fanelli and F. Piazza, *Analysis and forecast of covid-19 spreading in china, italy and*  
492 *france*, Chaos, Solitons & Fractals 134 (2020), p. 109761.
- 493 [15] J. Faraway and C. Chatfield, *Time series forecasting with neural networks: a comparative*  
494 *study using the air line data*, Journal of the Royal Statistical Society: Series C (Applied  
495 Statistics) 47 (1998), pp. 231–250.
- 496 [16] I. Ghosh and T. Chakraborty, *An integrated deterministic-stochastic approach for fore-*  
497 *casting the long-term trajectories of covid-19*, International Journal of Modeling, Simula-  
498 tion, and Scientific Computing (2021), p. 2141001.
- 499 [17] G. Grasselli, A. Pesenti, and M. Cecconi, *Critical care utilization for the covid-19 outbreak*  
500 *in lombardy, italy: early experience and forecast during an emergency response*, Jama 323  
501 (2020), pp. 1545–1546.
- 502 [18] G. Grasselli, A. Zangrillo, A. Zanella, M. Antonelli, L. Cabrini, A. Castelli, D. Cereda, A.

- 503 Coluccello, G. Foti, R. Fumagalli, *et al.*, *Baseline characteristics and outcomes of 1591*  
504 *patients infected with sars-cov-2 admitted to icus of the lombardy region, italy*, *Jama* 323  
505 (2020), pp. 1574–1581.
- 506 [19] W.j. Guan, Z.y. Ni, Y. Hu, W.h. Liang, C.q. Ou, J.x. He, L. Liu, H. Shan, C.l. Lei, D.S.  
507 Hui, *et al.*, *Clinical characteristics of coronavirus disease 2019 in china*, *New England*  
508 *journal of medicine* 382 (2020), pp. 1708–1720.
- 509 [20] T. Hastie, R. Tibshirani, and J. Friedman, *The elements of statistical learning: data*  
510 *mining, inference, and prediction*, Springer Science & Business Media, 2009.
- 511 [21] I. Holmdahl and C. Buckee, *Wrong but useful—what covid-19 epidemiologic models can*  
512 *and cannot tell us*, *New England Journal of Medicine* (2020).
- 513 [22] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu,  
514 *et al.*, *Clinical features of patients infected with 2019 novel coronavirus in wuhan, china*,  
515 *The lancet* 395 (2020), pp. 497–506.
- 516 [23] R. Hyndman, A.B. Koehler, J.K. Ord, and R.D. Snyder, *Forecasting with exponential*  
517 *smoothing: the state space approach*, Springer Science & Business Media, 2008.
- 518 [24] R.J. Hyndman and G. Athanasopoulos, *Forecasting: principles and practice*, OTexts, 2018.
- 519 [25] R.J. Hyndman, G. Athanasopoulos, C. Bergmeir, G. Caceres, L. Chhay, M. O’Hara-  
520 Wild, F. Petropoulos, S. Razbash, and E. Wang, *Package ‘forecast’*, Online] [https://cran.](https://cran.r-project.org/web/packages/forecast/forecast.pdf)  
521 [r-project. org/web/packages/forecast/forecast. pdf](https://cran.r-project.org/web/packages/forecast/forecast.pdf) (2020).
- 522 [26] R.J. Hyndman and B. Billah, *Unmasking the theta method*, *International Journal of Fore-*  
523 *casting* 19 (2003), pp. 287–290.
- 524 [27] R.J. Hyndman and A.B. Koehler, *Another look at measures of forecast accuracy*, *Inter-*  
525 *national journal of forecasting* 22 (2006), pp. 679–688.
- 526 [28] J.P. Ioannidis, S. Cripps, and M.A. Tanner, *Forecasting for covid-19 has failed*, *Interna-*  
527 *tional journal of forecasting* (2020).
- 528 [29] A.J. Kucharski, T.W. Russell, C. Diamond, Y. Liu, J. Edmunds, S. Funk, R.M. Eggo, F.  
529 Sun, M. Jit, J.D. Munday, *et al.*, *Early dynamics of transmission and control of covid-19:*  
530 *a mathematical modelling study*, *The lancet infectious diseases* (2020).
- 531 [30] L.I. Kuncheva, *Combining pattern classifiers: methods and algorithms*, John Wiley &  
532 Sons, 2004.
- 533 [31] Q. Li, W. Feng, and Y.H. Quan, *Trend and forecasting of the covid-19 outbreak in china*,  
534 *Journal of Infection* 80 (2020), pp. 469–496.
- 535 [32] A. Maleki, S. Nasser, M.S. Aminabad, and M. Hadi, *Comparison of arima and nnar*  
536 *models for forecasting water treatment plant’s influent characteristics*, *KSCE J. CIV.*  
537 *ENG.* 22 (2018), pp. 3233–3245.
- 538 [33] S.P. Meyn and R.L. Tweedie, *Markov chains and stochastic stability*, Springer Science &  
539 Business Media, 1993.
- 540 [34] L. Milačić, S. Jović, T. Vujović, and J. Miljković, *Application of artificial neural network*  
541 *with extreme learning machine for economic growth estimation*, *Physica A* 465 (2017), pp.  
542 285–288.
- 543 [35] A. Mosleh and G. Apostolakis, *The assessment of probability distributions from expert*  
544 *opinions with an application to seismic fragility curves*, *Risk analysis* 6 (1986), pp. 447–  
545 461.
- 546 [36] M.R. Oliveira and L. Torgo, *Ensembles for time series forecasting*, *J. Mach. Learn. Res.*  
547 39 (2014), pp. 360–370.
- 548 [37] W.H. Organization, *et al.*, *Who director-general’s opening remarks at the media briefing*  
549 *on covid-19-11 march 2020* (2020).
- 550 [38] A. Pasini, *Artificial neural networks for small dataset analysis*, *Journal of thoracic disease*  
551 7 (2015), p. 953.
- 552 [39] R.K. Paul, S. Samanta, M.R.K. Paul, and T. LazyData, *Package ‘waveletarima’*, *Seed*  
553 500 (2017), pp. 1–5.
- 554 [40] M. Perc, N. Gorišek Miksić, M. Slavinec, and A. Stožer, *Forecasting covid-19*, *Frontiers*  
555 *in Physics* 8 (2020), p. 127.
- 556 [41] D.B. Percival and A.T. Walden, *Wavelet methods for time series analysis*, Vol. 4, Cam-

- 557 bridge university press, 2000.
- 558 [42] D.B. Percival and A.T. Walden, *Spectral Analysis for Univariate Time Series*, Vol. 51,  
559 Cambridge University Press, 2020.
- 560 [43] F. Petropoulos and S. Makridakis, *Forecasting the novel coronavirus covid-19*, PloS one  
561 15 (2020), p. e0231236.
- 562 [44] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foun-  
563 dation for Statistical Computing, Vienna, Austria (2017). Available at <https://www.R-project.org/>.
- 564 [45] D. Ray, M. Salvatore, R. Bhattacharyya, L. Wang, J. Du, S. Mohammed, S. Purkayastha,  
565 A. Halder, A. Rix, D. Barker, *et al.*, *Predictions, role of interventions and effects of a*  
566 *historic national lockdown in india’s response to the covid-19 pandemic: data science call*  
567 *to arms*, Harvard data science review 2020 (2020).
- 568 [46] M.H.D.M. Ribeiro, R.G. da Silva, V.C. Mariani, and L. dos Santos Coelho, *Short-term*  
569 *forecasting covid-19 cumulative confirmed cases: Perspectives for brazil*, Chaos, Solitons  
570 & Fractals (2020), p. 109853.
- 571 [47] K. Roosa, Y. Lee, R. Luo, A. Kirpich, R. Rothenberg, J. Hyman, P. Yan, and G. Chowell,  
572 *Real-time forecasts of the covid-19 epidemic in china from february 5th to february 24th,*  
573 *2020*, Infectious Disease Modelling 5 (2020), pp. 256–263.
- 574 [48] L. Rosenbaum, *Facing covid-19 in italy—ethics, logistics, and therapeutics on the epi-*  
575 *demical’s front line*, New England Journal of Medicine 382 (2020), pp. 1873–1875.
- 576 [49] D.E. Rumelhart, G.E. Hinton, and R.J. Williams, *Learning internal representations by*  
577 *error propagation*, Tech. Rep., California Univ San Diego La Jolla Inst for Cognitive  
578 Science, 1985.
- 579 [50] F. Rustam, A.A. Reshi, A. Mehmood, S. Ullah, B.W. On, W. Aslam, and G.S. Choi,  
580 *Covid-19 future forecasting using supervised machine learning models*, IEEE access 8  
581 (2020), pp. 101489–101499.
- 582 [51] E. Spiliotis, V. Assimakopoulos, and S. Makridakis, *Generalizing the theta method for*  
583 *automatic forecasting*, European Journal of Operational Research 284 (2020), pp. 550–  
584 558.
- 585 [52] R. Sujath, J.M. Chatterjee, and A.E. Hassanien, *A machine learning forecasting model*  
586 *for covid-19 pandemic in india*, Stochastic Environmental Research and Risk Assessment  
587 (2020), p. 1.
- 588 [53] A. Trapletti, F. Leisch, and K. Hornik, *Stationary and integrated autoregressive neural*  
589 *network processes*, Neural Computation 12 (2000), pp. 2427–2450.
- 590 [54] A. Trilla, G. Trilla, and C. Daer, *The 1918 “spanish flu” in spain*, Clinical infectious  
591 diseases 47 (2008), pp. 668–673.
- 592 [55] R. Tweedie, *The existence of moments for stationary markov chains*, Journal of Applied  
593 Probability 20 (1983), pp. 191–196.
- 594 [56] J.T. Wu, K. Leung, and G.M. Leung, *Nowcasting and forecasting the potential domes-*  
595 *tic and international spread of the 2019-ncov outbreak originating in wuhan, china: a*  
596 *modelling study*, The Lancet 395 (2020), pp. 689–697.
- 597 [57] L. Wynants, B. Van Calster, M.M. Bonten, G.S. Collins, T.P. Debray, M. De Vos, M.C.  
598 Haller, G. Heinze, K.G. Moons, R.D. Riley, *et al.*, *Prediction models for diagnosis and*  
599 *prognosis of covid-19 infection: systematic review and critical appraisal*, bmj 369 (2020).
- 600 [58] G.P. Zhang, *Time series forecasting using a hybrid arima and neural network model*,  
601 Neurocomputing 50 (2003), pp. 159–175.
- 602 [59] G. Zhang, B.E. Patuwo, and M.Y. Hu, *Forecasting with artificial neural networks: The*  
603 *state of the art*, International journal of forecasting 14 (1998), pp. 35–62.
- 604 [60] Z. Zhuang, P. Cao, S. Zhao, Y. Lou, W. Wang, S. Yang, L. Yang, and D. He, *Estimation*  
605 *of local novel coronavirus (covid-19) cases in wuhan, china from off-site reported cases*  
606 *and population flow data from different sources*, medRxiv (2020).
- 607